

PCT

WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau



AN

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification ⁶ : C07K 1/00, 16/00, C07H 21/04, C12N 15/00		A1	(11) International Publication Number: WO 97/30070
			(43) International Publication Date: 21 August 1997 (21.08.97)
(21) International Application Number: PCT/US97/02318			Royersford, PA 19468 (US). WARD, Judith, M. [GB/GB]; Carmela Cottage, 19 Cotmandene, Dorking, Surrey RH4 2BT (GB). (74) Agents: GIMMI, Edward, R. et al.; SmithKline Beecham Corporation, Corporate Intellectual Property, UW2220, 709 Swedeland Road, P.O. Box 1539, King of Prussia, PA 19406-0939 (US). (81) Designated States: JP, US, European patent (AT, BE, CH, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE). Published <i>With international search report. Before the expiration of the time limit for amending the claims and to be republished in the event of the receipt of amendments.</i>
(22) International Filing Date: 19 February 1997 (19.02.97)			
(30) Priority Data: 60/011,888 20 February 1996 (20.02.96) US			
(71) Applicant (for all designated States except US): SMITHKLINE BEECHAM CORPORATION [US/US]; Corporate Intellectual Property, UW2220, 709 Swedeland Road, P.O. Box 1539, King of Prussia, PA 19406-0939 (US).			
(72) Inventors; and (75) Inventors/Applicants (for US only): BLACK, Michael, Terrence [GB/US]; 502 Milhouse Way, Chester Springs, PA 19425 (US). BURNHAM, Martin, Karl, Russel [US/US]; 2927 Tanglewood Lane, Norristown, PA 19403 (US). HODGSON, John, Edward [GB/GB]; 260 Lapp Road, Malvern, PA 19355 (GB). KNOWLES, David, Justin, Charles [GB/GB]; Downsview House, Redhill, Surrey RH1 6LY (GB). NICHOLAS, Richard, Oakley [GB/GB]; 355 Carmen Drive, Collegeville, PA 19426 (GB). PRATT, Julie, M. [GB/GB]; 77 Highfield Drive, Wigston, Leicester LE8 1NP (GB). REICHARD, Raymond, Winfield [US/US]; 2091 Clover Mill Road, Quakertown, PA 18951 (US). ROSENBERG, Martin [US/US]; 241 Mingo Road,			
(54) Title: NOVEL COMPOUNDS			
(57) Abstract This invention relates to newly identified Staphylococcal polynucleotides, polypeptides encoded by such polynucleotides, the uses of such polynucleotides and polypeptides, as well as the production of such polynucleotides and polypeptides and recombinant host cells transformed with the polynucleotides. This invention also relates to inhibiting the biosynthesis or action of such polynucleotides or polypeptides and to the use of such inhibitors in therapy.			

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AM	Armenia	GB	United Kingdom	MW	Malawi
AT	Austria	GE	Georgia	MX	Mexico
AU	Australia	GN	Guinea	NE	Niger
BB	Barbados	GR	Greece	NL	Netherlands
BE	Belgium	HU	Hungary	NO	Norway
BF	Burkina Faso	IE	Ireland	NZ	New Zealand
BG	Bulgaria	IT	Italy	PL	Poland
BJ	Benin	JP	Japan	PT	Portugal
BR	Brazil	KE	Kenya	RO	Romania
BY	Belarus	KG	Kyrgyzstan	RU	Russian Federation
CA	Canada	KP	Democratic People's Republic of Korea	SD	Sudan
CF	Central African Republic	KR	Republic of Korea	SE	Sweden
CG	Congo	KZ	Kazakhstan	SG	Singapore
CH	Switzerland	LJ	Liechtenstein	SI	Slovenia
CI	Côte d'Ivoire	LK	Sri Lanka	SK	Slovakia
CM	Cameroon	LR	Liberia	SN	Senegal
CN	China	LT	Lithuania	SZ	Swaziland
CS	Czechoslovakia	LU	Luxembourg	TD	Chad
CZ	Czech Republic	LV	Latvia	TG	Togo
DE	Germany	MC	Monaco	TJ	Tajikistan
DK	Denmark	MD	Republic of Moldova	TT	Trinidad and Tobago
EE	Estonia	MG	Madagascar	UA	Ukraine
ES	Spain	ML	Mali	UG	Uganda
FI	Finland	MN	Mongolia	US	United States of America
FR	France	MR	Mauritania	UZ	Uzbekistan
GA	Gabon			VN	Viet Nam

NOVEL COMPOUNDS

Field of the Invention:

This invention relates to newly identified polynucleotides, particularly Staphylococcal polynucleotides, polypeptides encoded by such polynucleotides, the uses of such polynucleotides and polypeptides, as well as the production of such polynucleotides and polypeptides and recombinant host cells transformed with the polynucleotides. This invention also relates to activating or inhibiting the biosynthesis or action of such polynucleotides or polypeptides and to the use of such activators or inhibitors in therapy.

Background of the Invention:

The Staphylococci make up a medically important genera of microbes. They are known to produce two types of disease, invasive and toxigenic. Invasive infections are characterized generally by abscess formation effecting both skin surfaces and deep tissues. *Staphylococcus aureus* (*S. aureus*) is the second leading cause of bacteremia in cancer patients. Osteomyelitis, septic arthritis, septic thrombophlebitis and acute bacterial endocarditis are also relatively common. There are at least three clinical conditions resulting from the toxigenic properties of Staphylococci. The manifestation of these diseases result from the actions of exotoxins as opposed to tissue invasion and bacteremia. These conditions include: Staphylococcal food poisoning, scalded skin syndrome and toxic shock syndrome.

While certain Staphylococcal proteins associated with pathogenicity have been identified, e.g., coagulase, hemolysins, leucocidins and exo and enterotoxins, very little is known concerning the temporal expression of such genes during infection and disease progression in a mammalian host. Discovering the sets of genes the bacterium is likely to be expressing at the different stages of infection, particularly when an infection is established, provides critical information for the screening and characterization of novel antibacterials which can interrupt pathogenesis. In addition to providing a fuller understanding of known proteins, such an approach will identify previously unrecognized targets.

Brief Description of the Invention:

This invention provides novel protein from *Staphylococcus aureus* WCUH29, characterized in that it comprises the amino acid sequence given herein or a fragment, analogue or derivative thereof.

Also provided by the invention is an isolated polypeptide comprising an acid sequence selected from the group consisting of SEQ ID NO:1 through SEQ ID NO:623, or any combination of these SEQ ID NOS (SEQ ID NOS:1 through 623).

Further provided by the invention is an isolated nucleic acid (herein also "polynucleotide") encoding one of the amino acid sequences of SEQ ID NO:1 through SEQ ID NO:623, or an isolated polynucleotide sequence selected from the group consisting of SEQ ID NO:624 through 1165, or any combination of these SEQ ID NOS (SEQ ID NOS:624 through 1165), or any polynucleotide sequences capable of hybridizing therewith under stringent conditions.

In accordance with another aspect of the present invention, there are provided polynucleotides (DNA or RNA) which encode such polypeptides.

In particular the invention provides polynucleotides having the DNA sequences given herein.

The invention also relates to novel oligonucleotides derived from the sequences given herein which can act as PCR primers in the process herein described to determine whether or not the gene identified herein, particularly a *Staphylococcus aureus* gene, in whole or in part is expressed in infected tissue. It is recognized that such sequences will also have utility in diagnosis of the stage of infection and type of infection the pathogen has attained. The proteins so identified are also useful as targets in screens designed to identify antimicrobial compounds.

Detailed Description of the Invention:

Each of the polynucleotide sequences provided herein, particularly the DNA sequences, may be used in the discovery and development of antibacterial compounds. Because each of the sequences contains an open reading frame (ORF) or putative gene with an appropriate initiation and termination codons, the encoded protein upon expression can be used as a target for the screening of antimicrobial compounds, such as drugs. Additionally, the DNA sequences encoding the amino terminal regions of the encoded protein can be used to construct antisense or ribozyme sequences to control the expression of the coding sequence of interest. Furthermore, many of the sequences disclosed herein also provide regions upstream and downstream from the encoding sequence. These sequences are useful as a source of regulatory elements for the control of bacterial gene expression. Such sequences are conveniently isolated by restriction enzyme action or synthesized chemically and introduced, for example, into promoter identification strains. These strains contain a reporter structural gene sequence located downstream from a restriction site such that if an active promoter is inserted, the reporter gene will be expressed.

Although each of the sequences may be employed as described above, this invention also provides several means for identifying particularly useful target genes. The first of these approaches entails searching appropriate databases for sequence matches in related organisms. Thus, if a homologue exists, the Staphylococcal-like form of this gene would likely play an analogous role. For example, a Staphylococcal protein identified as homologous to a cell surface protein in another organism would be useful as a vaccine candidate. To the extent such homologies have been identified for the sequences disclosed herein they are reported along with the coding sequence.

Recently techniques have become available to evaluate temporal gene expression in bacteria, particularly as it applies to viability under laboratory and infection conditions. A number of methods can be used to identify genes which are essential to survival *per se*, or essential to the establishment/maintenance of an infection. Identification of an ORF unknown by one of these methods yields additional information about its function and permits the selection of such an ORF for further development as a screening target. Briefly, these approaches include:

1) Signature Tagged Mutagenesis (STM)

This technique is described by Hensel *et al.*, Science 269: 400-403(1995), the contents of which is incorporated by reference for background purposes.

Signature tagged mutagenesis identifies genes necessary for the establishment/maintenance of infection in a given infection model.

The basis of the technique is the random mutagenesis of target organism by various means (e.g., transposons) such that unique DNA sequence tags are inserted in close proximity to the site of mutation. The tags from a mixed population of bacterial mutants and bacteria recovered from an infected hosts are detected by amplification, radiolabeling and hybridization analysis. Mutants attenuated in virulence are revealed by absence of the tag from the pool of bacteria recovered from infected hosts.

2) In Vivo Expression Technology (IVET)

This technique is described by Camilli *et al.*, Proc. Nat'l. Acad. Sci. USA. 91:2634-2638 (1994), the contents of which is incorporated by reference for background purposes.

IVET identifies genes up-regulated during infection when compared to laboratory cultivation, implying an important role in infection. ORF identified by this technique are implied to have a significant role in infection establishment/maintenance.

In this technique random chromosomal fragments of target organism are cloned upstream of a promoter-less recombinase gene in a plasmid vector. This construct is

introduced into the target organism which carries an antibiotic resistance gene flanked by resolvase sites. Growth in the presence of the antibiotic removes from the population those fragments cloned into the plasmid vector capable of supporting transcription of the recombinase gene and therefore have caused loss of antibiotic resistance. The resistant pool is introduced into a host and at various times after infection bacteria may be recovered and assessed for the presence of antibiotic resistance. The chromosomal fragment carried by each antibiotic sensitive bacterium should carry a promoter or portion of a gene normally upregulated during infection. Sequencing upstream of the recombinase gene allows identification of the up regulated gene.

3) Differential display

This technique is described by Chuang *et al.*, *J. Bacteriol.* 175:2026-2036 (1993), the contents of which is incorporated by reference for background purposes. This method identifies those genes which are expressed in an organism by identifying mRNA present using randomly-primed RT-PCR. By comparing pre-infection and post infection profiles, genes up and down regulated during infection can be identified and the RT-PCR product sequenced and matched to ORF 'unknowns'.

4) Generation of conditional lethal mutants by transposon mutagenesis.

This technique, described by de Lorenzo, V. *et al.*, *Gene* 123:17-24 (1993); Neuwald, A. F. *et al.*, *Gene* 125: 69-73(1993); and Takiff, H. E. *et al.*, *J. Bacteriol.* 174:1544-1553(1992), the contents of which is incorporated by reference for background purposes, identifies genes whose expression are essential for cell viability.

In this technique transposons carrying controllable promoters, which provide transcription outward from the transposon in one or both directions, are generated. Random insertion of these transposons into target organisms and subsequent isolation of insertion mutants in the presence of inducer of promoter activity ensures that insertions which separate promoter from coding region of a gene whose expression is essential for cell viability will be recovered. Subsequent replica plating in the absence of inducer identifies such insertions, since they fail to survive. Sequencing of the flanking regions of the transposon allows identification of site of insertion and identification of the gene disrupted. Close monitoring of the changes in cellular processes/morphology during growth in the absence of inducer yields information on likely function of the gene. Such monitoring could include flow cytometry (cell division, lysis, redox potential, DNA replication), incorporation of radiochemically labeled precursors into DNA, RNA, protein, lipid,

peptidoglycan, monitoring reporter enzyme gene fusions which respond to known cellular stresses.

5) Generation of conditional lethal mutants by chemical mutagenesis.

This technique is described by Beckwith, J. Methods in Enzymology 204:

3-18(1991), the contents of which are incorporated herein by reference for background purposes. In this technique random chemical mutagenesis of target organism, growth at temperature other than physiological temperature (permissive temperature) and subsequent replica plating and growth at different temperature (e.g. 42°C to identify ts, 25°C to identify cs) are used to identify those isolates which now fail to grow (conditional mutants). As above close monitoring of the changes upon growth at the non-permissive temperature yields information on the function of the mutated gene. Complementation of conditional lethal mutation by library from target organism and sequencing of complementing gene allows matching with unknown ORF.

Each of these techniques may have advantages or disadvantage depending on the particular application. The skilled artisan would choose the approach that is the most relevant with the particular end use in mind. For example, some genes might be recognized as essential for infection but in reality are only necessary for the initiation of infection and so their products would represent relatively unattractive targets for antibacterials developed to cure established and chronic infections.

In yet another variation, a suitably labelled oligonucleotide probe which anneals specifically to the bacterial ribosomal RNA in Northern blots of bacterial RNA preparations from infected tissue is employed. Using the more abundant ribosomal RNA as a hybridization target greatly facilitates the optimization of a protocol to purify bacterial RNA of a suitable size for RT-PCR from infected tissue.

Use of the of these technologies when applied to the ORFs of the present invention enables identification of bacterial proteins expressed during infection, inhibitors of which would have utility in anti-bacterial therapy.

S. aureus WCUH 29 has been deposited at the National Collection of Industrial and Marine Bacteria Ltd. (NCIMB), Aberdeen, Scotland under number NCIMB 40771 on 11 September 1995.

The nucleotide sequences disclosed herein can be obtained by synthetic chemical techniques known in the art or can be obtained from *S. aureus* WCUH 29 by probing a DNA preparation with probes constructed from the particular sequences disclosed herein. Alternatively, oligonucleotides derived from a disclosed sequence can act as PCR primers

in a process of PCR-based cloning of the sequence from a bacterial genomic source. It is recognized that such sequences will also have utility in diagnosis of the stage of infection and type of infection the pathogen has attained.

To obtain the polynucleotide encoding the protein using the DNA sequence given
5 herein typically a library of clones of chromosomal DNA of *S. aureus* WCUH29 in *E. coli*
or some other suitable host is probed with a radiolabelled oligonucleotide, preferably a
17mer or longer, derived from the partial sequence. Clones carrying DNA identical to that
of the probe can then be distinguished using high stringency washes. By sequencing the
individual clones thus identified with sequencing primers designed from the original
10 sequence it is then possible to extend the sequence in both directions to determine the full
gene sequence. Conveniently such sequencing is performed using denatured double
stranded DNA prepared from a plasmid clone. Suitable techniques are described by
Maniatis, T., Fritsch, E.F. and Sambrook, J. in MOLECULAR CLONING, A Laboratory
Manual, 2nd edition, 1989, Cold Spring Harbor Laboratory (see: Screening By
15 Hybridization 1.90 and Sequencing Denatured Double-Stranded DNA Templates 13.70).

A polynucleotide of the present invention may be in the form of RNA or in the
form of DNA, which DNA includes cDNA, genomic DNA, and synthetic DNA. The DNA
may be double-stranded or single-stranded, and if single stranded may be the coding strand
or non-coding (anti-sense) strand. The coding sequence which encodes the polypeptide
20 may be identical to the coding sequence shown or may be a different coding sequence
which coding sequence, as a result of the redundancy or degeneracy of the genetic code,
encoding the same polypeptide.

The present invention includes variants of the hereinabove described
polynucleotides which encode fragments, analogues and derivatives of the polypeptide
25 characterized by the deduced amino acid sequence given herein. The variant of the
polynucleotide may be a naturally occurring allelic variant of the polynucleotide or a non-
naturally occurring variant of the polynucleotide. In addition to the standard A, G, C, T/U
representations for nucleic acid bases, the term "N" is also used in certain polynucleotides
of the invention. "N" means that any of the four DNA or RNA bases may appear at such a
30 designated position in the DNA or RNA sequence, except that in preferred embodiments N
cannot be a base that when taken in combination with adjacent nucleotide positions, when
read in the correct reading frame, would have the effect of generating a premature
termination codon in such reading frame.

Thus, the present invention includes polynucleotides encoding the same polypeptide characterized by the deduced amino acid sequence given herein as well as variants of such polynucleotides which variants encode for a fragment, derivative or analogue of the polypeptide. Such nucleotide variants include deletion variants, substitution variants and addition or insertion variants.

The polynucleotide may have a coding sequence which is a naturally occurring allelic variant of the coding sequence characterized by the DNA sequence disclosed herein. As known in the art, an allelic variant is an alternate form of a polynucleotide sequence which may have a substitution, deletion or addition of one or more nucleotides, which does not substantially alter the function of the encoded polypeptide.

The polynucleotide which encodes for the mature polypeptide, may include only the coding sequence for the mature polypeptide or the coding sequence for the mature polypeptide and additional coding sequence such as a leader or secretory sequence or a proprotein sequence.

Thus, the term "polynucleotide encoding a polypeptide" encompasses a polynucleotide which includes only coding sequence for the polypeptide as well as a polynucleotide which includes additional coding and/or non-coding sequence.

The present invention therefore includes polynucleotides, wherein the coding sequence for the mature polypeptide may be fused in the same reading frame to a polynucleotide sequence which aids in expression and secretion of a polypeptide from a host cell, for example, a leader sequence which functions as a secretory sequence for controlling transport of a polypeptide from the cell. The polypeptide having a leader sequence is a preprotein and may have the leader sequence cleaved by the host cell to form the mature form of the polypeptide. The polynucleotides may also encode for a proprotein which is the mature protein plus additional 5' amino acid residues. A mature protein having a prosequence is a proprotein and is an inactive form of the protein. Once the prosequence is cleaved an active mature protein remains.

Thus, for example, the polynucleotide of the present invention may encode for a mature protein, or for a protein having a prosequence or for a protein having both a prosequence and a presequence (leader sequence). Further, the amino acid sequences provided herein show a methionine residue at the NH₂-terminus. It is appreciated, however, that during post-translational modification of the peptide, this residue may be deleted. Accordingly, this invention contemplates the use of both the methionine-containing and the methionineless amino terminal variants of each protein disclosed herein.

The polynucleotides of the present invention may also have the coding sequence fused in frame to a marker sequence at either the 5' or 3' terminus of the gene which allows for purification of the polypeptide of the present invention. The marker sequence may be a hexa-histidine tag supplied by the pQE series of vectors (supplied commercially by Quiagen Inc.) to provide for purification of the polypeptide fused to the marker in the case of a bacterial host.

The present invention further relates to polynucleotides which hybridize to the hereinabove-described sequences if there is at least 50% or 60% and preferably at least 70%, 80% or 90% identity between the sequences. The present invention particularly relates to Staphylococcal polynucleotides which hybridize under stringent conditions to the hereinabove-described polynucleotides. As herein used, the term "stringent conditions" means hybridization will occur only if there is at least 95% and preferably at least 97% identity between the sequences. The polynucleotides which hybridize to the hereinabove described polynucleotides in a preferred embodiment encode polypeptides which retain substantially the same biological function or activity as the polypeptide characterized by the deduced amino acid sequence given herein.

The deposit referred to herein will be maintained under the terms of the Budapest Treaty on the International Recognition of the Deposit of Micro-organisms for purposes of Patent Procedure. These deposits are provided merely as convenience to those of skill in the art and are not an admission that a deposit is required under 35 U.S.C. §112. The sequence of the polynucleotides contained in the deposited material, as well as the amino acid sequence of the polypeptides encoded thereby, are incorporated herein by reference and are controlling in the event of any conflict with any description of sequences herein. A license may be required to make, use or sell the deposited material, and no such license is hereby granted.

The terms "fragment," "derivative" and "analogue" when referring to the polypeptide characterized by the deduced amino acid sequence herein, means a polypeptide which retains essentially the same biological function or activity as such polypeptide. Thus, an analogue includes a proprotein which can be activated by cleavage of the proprotein portion to produce an active mature polypeptide.

The polypeptide of the present invention may be a recombinant polypeptide, a natural polypeptide or a synthetic polypeptide, preferably a recombinant polypeptide.

The fragment, derivative or analogue of the polypeptide characterized by the deduced amino acid sequence herein may be (i) one in which one or more of the amino acid

residues are substituted with a conserved or non-conserved amino acid residue (preferably a conserved amino acid residue) and such substituted amino acid residue may or may not be one encoded by the genetic code, or (ii) one in which one or more of the amino acid residues includes a substituent group, or (iii) one in which the polypeptide is fused with
5 another compound, such as a compound to increase the half-life of the polypeptide (for example, polyethylene glycol), or (iv) one in which the additional amino acids are fused to the polypeptide, such as a leader or secretory sequence or a sequence which is employed for purification of the polypeptide or a proprotein sequence. Such fragments, derivatives and analogues are deemed to be within the scope of those skilled in the art from the teachings
10 herein.

The polypeptides and polynucleotides of the present invention are preferably provided in an isolated form, and preferably are purified to homogeneity.

The term "isolated" means that the material is removed from its original environment (e.g., the natural environment if it is naturally occurring). For example, a
15 naturally-occurring polynucleotide or polypeptide present in a living animal is not isolated, but the same polynucleotide or polypeptide, separated from some or all of the coexisting materials in the natural system, is isolated. Such polynucleotides could be part of a vector and/or such polynucleotides or polypeptides could be part of a composition, and still be isolated in that such vector or composition is not part of its natural environment.

20 The present invention also relates to vectors which include polynucleotides of the present invention, host cells which are genetically engineered with vectors of the invention and the production of polypeptides of the invention by recombinant techniques.

In accordance with yet a further aspect of the present invention, there is therefore provided a process for producing the polypeptide of the invention by recombinant
25 techniques by expressing a polynucleotide encoding said polypeptide in a host and recovering the expressed product. Alternatively, the polypeptides of the invention can be synthetically produced by conventional peptide synthesizers.

Host cells are genetically engineered (transduced or transformed or transfected) with the vectors of this invention which may be, for example, a cloning vector or an
30 expression vector. The vector may be, for example, in the form of a plasmid, a cosmid, a phage, etc. The engineered host cells can be cultured in conventional nutrient media modified as appropriate for activating promoters, selecting transformants or amplifying the genes. The culture conditions, such as temperature, pH and the like, are those previously

used with the host cell selected for expression, and will be apparent to the ordinarily skilled artisan.

Suitable expression vectors include chromosomal, nonchromosomal and synthetic DNA sequences, e.g., bacterial plasmids; phage DNA; baculovirus; yeast plasmids; vectors
5 derived from combinations of plasmids and phage DNA. However, any other vector may be used as long as it is replicable and viable in the host.

The appropriate DNA sequence may be inserted into the vector by a variety of procedures. In general, the DNA sequence is inserted into an appropriate restriction endonuclease site(s) by procedures known in the art.

10 The DNA sequence in the expression vector is operatively linked to an appropriate expression control sequence(s) (promoter) to direct mRNA synthesis. As representative examples of such promoters, there may be mentioned: LTR or SV40 promoter, the *E. coli* *lac* or *trp*, the phage lambda P_L promoter and other promoters known to control expression of genes in eukaryotic or prokaryotic cells or their viruses. The expression vector also
15 contains a ribosome binding site for translation initiation and a transcription terminator. The vector may also include appropriate sequences for amplifying expression.

In addition, the expression vectors preferably contain one or more selectable marker genes to provide a phenotypic trait for selection of transformed host cells such as dihydrofolate reductase or neomycin resistance for eukaryotic cell culture, or such as
20 tetracycline or ampicillin resistance in *E. coli*.

The gene can be placed under the control of a promoter, ribosome binding site (for bacterial expression) and, optionally, an operator (collectively referred to herein as "control" elements), so that the DNA sequence encoding the desired protein is transcribed into RNA in the host cell transformed by a vector containing this expression construction. The coding
25 sequence may or may not contain a signal peptide or leader sequence. The polypeptides of the present invention can be expressed using, for example, the *E. coli* *tac* promoter or the protein A gene (*spa*) promoter and signal sequence. Leader sequences can be removed by the bacterial host in post-translational processing. See, e.g., U.S. Patent Nos. 4,431,739; 4,425,437; 4,338,397. Promoter regions can be selected from any desired gene using CAT
30 (chloramphenicol transferase) vectors or other vectors with selectable markers. Two appropriate vectors are PKK232-8 and PCM7. Particular named bacterial promoters include *lacI*, *lacZ*, T3, T7, *gpt*, lambda P_R, P_L and *trp*. Eukaryotic promoters include CMV immediate early, HSV thymidine kinase, early and late SV40, LTRs from retrovirus, and

mouse metallothionein-I. Selection of the appropriate vector and promoter is well within the level of ordinary skill in the art.

5 In addition to control sequences, it may be desirable to add regulatory sequences which allow for regulation of the expression of the protein sequences relative to the growth of the host cell. Regulatory sequences are known to those of skill in the art, and examples include those which cause the expression of a gene to be turned on or off in response to a chemical or physical stimulus, including the presence of a regulatory compound. Other types of regulatory elements may also be present in the vector, for example, enhancer sequences.

10 An expression vector is constructed so that the particular coding sequence is located in the vector with the appropriate regulatory sequences, the positioning and orientation of the coding sequence with respect to the control sequences being such that the coding sequence is transcribed under the "control" of the control sequences (i.e., RNA polymerase which binds to the DNA molecule at the control sequences transcribes the coding
15 sequence). Modification of the coding sequences may be desirable to achieve this end. For example, in some cases it may be necessary to modify the sequence so that it may be attached to the control sequences with the appropriate orientation; i.e., to maintain the reading frame. The control sequences and other regulatory sequences may be ligated to the coding sequence prior to insertion into a vector, such as the cloning vectors described
20 above. Alternatively, the coding sequence can be cloned directly into an expression vector which already contains the control sequences and an appropriate restriction site.

Generally, recombinant expression vectors will include origins of replication and selectable markers permitting transformation of the host cell, e.g., the ampicillin resistance gene of *E. coli* and *S. cerevisiae* TRP1 gene, and a promoter derived from a highly-
25 expressed gene to direct transcription of a downstream structural sequence. The heterologous structural sequence is assembled in appropriate phase with translation initiation and termination sequences, and preferably, a leader sequence capable of directing secretion of translated protein into the periplasmic space or extracellular medium. Optionally, the heterologous sequence can encode a fusion protein including an N-terminal
30 identification peptide imparting desired characteristics, e.g., stabilization or simplified purification of expressed recombinant product.

The vector containing the appropriate DNA sequence as hereinabove described, as well as an appropriate promoter or control sequence, may be employed to transform an appropriate host to permit the host to express the protein.

More particularly, the present invention also includes recombinant constructs comprising one or more of the sequences as broadly described above. The constructs comprise a vector, such as a plasmid or viral vector, into which a sequence of the invention has been inserted, in a forward or reverse orientation. In a preferred aspect of this embodiment, the construct further comprises regulatory sequences, including, for example, a promoter, operably linked to the sequence. Large numbers of suitable vectors and promoters are known to those of skill in the art, and are commercially available. The following vectors are provided by way of example. Bacterial: pET-3 vectors (Stratagene), pQE70, pQE60, pQE-9 (Qiagen), pbs, pD10, phagescript, psiX174, pbluescript SK, pbsks, pNH8A, pNH16a, pNH18A, pNH46A (Stratagene); ptrc99a, pKK223-3, pKK233-3, pDR540, pRIT5 (Pharmacia). Eukaryotic: pBlueBacIII (Invitrogen), pWLNEO, pSV2CAT, pOG44, pXT1, pSG (Stratagene) pSVK3, pBPV, pMSG, pSVL (Pharmacia). However, any other plasmid or vector may be used as long as they are replicable and viable in the host.

Examples of recombinant DNA vectors for cloning and host cells which they can transform include the bacteriophage λ (*E. coli*), pBR322 (*E. coli*), pACYC177 (*E. coli*), pKT230 (gram-negative bacteria), pGV1106 (gram-negative bacteria), pLAFR1 (gram-negative bacteria), pME290 (non-*E. coli* gram-negative bacteria), pHV14 (*E. coli* and *Bacillus subtilis*), pBD9 (*Bacillus*), pIJ61 (*Streptomyces*), pUC6 (*Streptomyces*), YIp5 (*Saccharomyces*), a baculovirus insect cell system, YCp19 (*Saccharomyces*). See, generally, "DNA Cloning": Vols. I & II, Glover *et al.* ed. IRL Press Oxford (1985) (1987) and; T. Maniatis *et al.* ("Molecular Cloning" Cold Spring Harbor Laboratory (1982).

In some cases, it may be desirable to add sequences which cause the secretion of the polypeptide from the host organism, with subsequent cleavage of the secretory signal.

Polypeptides can be expressed in host cells under the control of appropriate promoters. Cell-free translation systems can also be employed to produce such proteins using RNAs derived from the DNA constructs of the present invention. Appropriate cloning and expression vectors for use with prokaryotic and eukaryotic hosts are described by Sambrook, *et al.*, Molecular Cloning: A Laboratory Manual, Second Edition, Cold Spring Harbor, N.Y., (1989), the disclosure of which is hereby incorporated by reference.

Following transformation of a suitable host strain and growth of the host strain to an appropriate cell density, the selected promoter is induced by appropriate means (e.g., temperature shift or chemical induction) and cells are cultured for an additional period.

Cells are typically harvested by centrifugation, disrupted by physical or chemical means, and the resulting crude extract retained for further purification.

Microbial cells employed in expression of proteins can be disrupted by any convenient method, including freeze-thaw cycling, sonication, mechanical disruption, or use
5 of cell lysing agents, such methods are well known to those skilled in the art.

Depending on the expression system and host selected, the polypeptide of the present invention may be produced by growing host cells transformed by an expression vector described above under conditions whereby the polypeptide of interest is expressed. The polypeptide is then isolated from the host cells and purified. If the expression system
10 secretes the polypeptide into growth media, the polypeptide can be purified directly from the media. If the polypeptide is not secreted, it is isolated from cell lysates or recovered from the cell membrane fraction. Where the polypeptide is localized to the cell surface, whole cells or isolated membranes can be used as an assayable source of the desired gene product. Polypeptide expressed in bacterial hosts such as *E. coli* may require isolation from
15 inclusion bodies and refolding. Where the mature protein has a very hydrophobic region which leads to an insoluble product of overexpression, it may be desirable to express a truncated protein in which the hydrophobic region has been deleted. The selection of the appropriate growth conditions and recovery methods are within the skill of the art.

The polypeptide can be recovered and purified from recombinant cell cultures by
20 methods including ammonium sulphate or ethanol precipitation, acid extraction, anion or cation exchange chromatography, phosphocellulose chromatography, hydrophobic interaction chromatography, affinity chromatography, hydroxylapatite chromatography and lectin chromatography. Protein refolding steps can be used, as necessary, in completing configuration of the mature protein. Finally, high performance liquid chromatography
25 (HPLC) can be employed for final purification steps.

Depending upon the host employed in a recombinant production procedure, the polypeptides of the present invention may be glycosylated or may be non-glycosylated. Polypeptides of the invention may also include an initial methionine amino acid residue.

A "replicon" is any genetic element (e.g., plasmid, chromosome, virus) that
30 functions as an autonomous unit of DNA replication *in vivo*; i.e., capable of replication under its own control.

A "vector" is a replicon, such as a plasmid, phage, or cosmid, to which another DNA segment may be attached so as to bring about the replication of the attached segment.

A "double-stranded DNA molecule" refers to the polymeric form of deoxyribonucleotides (bases adenine, guanine, thymine, or cytosine) in a double-stranded helix, both relaxed and supercoiled. This term refers only to the primary and secondary structure of the molecule, and does not limit it to any particular tertiary forms. Thus, this term includes double-stranded DNA found, *inter alia*, in linear DNA molecules (e.g., restriction fragments), viruses, plasmids, and chromosomes. In discussing the structure of particular double-stranded DNA molecules, sequences may be described herein according to the normal convention of giving only the sequence in the 5' to 3' direction along the nontranscribed strand of DNA (i.e., the strand having the sequence homologous to the mRNA).

A DNA "coding sequence of" or a "nucleotide sequence encoding" a particular protein, is a DNA sequence which is transcribed and translated into a polypeptide when placed under the control of appropriate regulatory sequences.

A "promoter sequence" is a DNA regulatory region capable of binding RNA polymerase in a cell and initiating transcription of a downstream (3' direction) coding sequence. For purposes of defining the present invention, the promoter sequence is bound at the 3' terminus by a translation start codon (e.g., ATG, GTG) of a coding sequence and extends upstream (5' direction) to include the minimum number of bases or elements necessary to initiate transcription at levels detectable above background. Within the promoter sequence will be found a transcription initiation site (conveniently defined by mapping with nuclease S1), as well as protein binding domains (consensus sequences) responsible for the binding of RNA polymerase. Eukaryotic promoters will often, but not always, contain "TATA" boxes and "CAT" boxes. Prokaryotic promoters contain Shine-Dalgarno sequences in addition to the -10 and -35 consensus sequences.

DNA "control sequences" refers collectively to promoter sequences, ribosome binding sites, polyadenylation signals, transcription termination sequences, upstream regulatory domains, enhancers, and the like, which collectively provide for the expression (i.e., the transcription and translation) of a coding sequence in a host cell.

A control sequence "directs the expression" of a coding sequence in a cell when RNA polymerase will bind the promoter sequence and transcribe the coding sequence into mRNA, which is then translated into the polypeptide encoded by the coding sequence.

A "host cell" is a cell which has been transformed or transfected, or is capable of transformation or transfection by an exogenous DNA sequence.

A cell has been "transformed" by exogenous DNA when such exogenous DNA has been introduced inside the cell membrane. Exogenous DNA may or may not be integrated (covalently linked) into chromosomal DNA making up the genome of the cell. In prokaryotes and yeasts, for example, the exogenous DNA may be maintained on an episomal element, such as a plasmid. With respect to eukaryotic cells, a stably transformed or transfected cell is one in which the exogenous DNA has become integrated into the chromosome so that it is inherited by daughter cells through chromosome replication. This stability is demonstrated by the ability of the eukaryotic cell to establish cell lines or clones comprised of a population of daughter cell containing the exogenous DNA.

A "clone" is a population of cells derived from a single cell or common ancestor by mitosis. A "cell line" is a clone of a primary cell that is capable of stable growth *in vitro* for many generations.

A "heterologous" region of a DNA construct is an identifiable segment of DNA within or attached to another DNA molecule that is not found in association with the other molecule in nature.

In accordance with yet a further aspect of the present invention, there is provided the use of a polypeptide of the invention for therapeutic or prophylactic purposes, for example, as an antibacterial agent or a vaccine.

In accordance with another aspect of the present invention, there is provided the use of a polynucleotide of the invention for therapeutic or prophylactic purposes, in particular genetic immunization.

In accordance with yet another aspect of the present invention, there are provided inhibitors to such polypeptides, useful as antibacterial agents. In particular, there are provided antibodies against such polypeptides.

Another aspect of the invention is a pharmaceutical composition comprising the above polypeptide, polynucleotide or inhibitor of the invention and a pharmaceutically acceptable carrier.

In a particular aspect the invention provides the use of an inhibitor of the invention as an antibacterial agent.

The invention further relates to the manufacture of a medicament for such uses.

The polypeptide may be used as an antigen for vaccination of a host to produce specific antibodies which have anti-bacterial action.

The polypeptides or cells expressing them can be used as an immunogen to produce antibodies thereto. These antibodies can be, for example, polyclonal or monoclonal

antibodies. The term antibodies also includes chimeric, single chain, and humanized or simianized antibodies, as well as Fab fragments, or the product of an Fab expression library. Various procedures known in the art may be used for the production of such antibodies and fragments.

- 5 Antibodies generated against the polypeptides of the present invention can be obtained by direct injection of the polypeptides into an animal or by administering the polypeptides to an animal, preferably a nonhuman. The antibody so obtained will then bind the polypeptides itself. In this manner, even a sequence encoding only a fragment of the polypeptides can be used to generate antibodies binding the whole native polypeptides.
- 10 Such antibodies can then be used to isolate the polypeptide from tissue expressing that polypeptide.

Polypeptide derivatives include antigenically or immunologically equivalent derivatives which form a particular aspect of this invention.

- 15 The term 'antigenically equivalent derivative' as used herein encompasses a polypeptide or its equivalent which will be specifically recognized by certain antibodies which, when raised to the protein or polypeptide according to the present invention, interfere with the interaction between pathogen and mammalian host.

- 20 The term 'immunologically equivalent derivative' as used herein encompasses a peptide or its equivalent which when used in a suitable formulation to raise antibodies in a vertebrate, the antibodies act to interfere with the interaction between pathogen and mammalian host.

- In particular derivatives which are slightly longer or slightly shorter than the native protein or polypeptide fragment of the present invention may be used. In addition, polypeptides in which one or more of the amino acid residues are modified may be used.
- 25 Such peptides may, for example, be prepared by substitution, addition, or rearrangement of amino acids or by chemical modification thereof. All such substitutions and modifications are generally well known to those skilled in the art of peptide chemistry.

- The polypeptide, such as an antigenically or immunologically equivalent derivative or a fusion protein thereof is used as an antigen to immunize a mouse or other animal such as a rat or chicken. The fusion protein may provide stability to the polypeptide. The antigen may be associated, for example by conjugation, with an immunogenic carrier protein for example bovine serum albumin (BSA) or keyhole limpet haemocyanin (KLH). Alternatively a multiple antigenic peptide comprising multiple copies of the protein or
- 30

polypeptide, or an antigenically or immunologically equivalent polypeptide thereof may be sufficiently antigenic to improve immunogenicity so as to obviate the use of a carrier.

For preparation of monoclonal antibodies, any technique which provides antibodies produced by continuous cell line cultures can be used. Examples include the hybridoma technique (Kohler and Milstein, Nature, 256:495-497(1975)), the trioma technique, the human B-cell hybridoma technique (Kozbor et al., Immunology Today 4:72(1983)), and the EBV-hybridoma technique to produce human monoclonal antibodies (Cole, et al., 1985, in Monoclonal Antibodies and Cancer Therapy, Alan R. Liss, Inc., pp. 77-96).

Techniques described for the production of single chain antibodies (U.S. Patent 4,946,778) can be adapted to produce single chain antibodies to immunogenic polypeptide products of this invention.

Using the procedure of Kohler and Milstein (supra, (1975)), antibody-containing cells from the immunized mammal are fused with myeloma cells to create hybridoma cells secreting monoclonal antibodies.

The hybridomas are screened to select a cell line with high binding affinity and favorable cross reaction with other staphylococcal species using one or more of the original polypeptide and/or the fusion protein. The selected cell line is cultured to obtain the desired Mab.

Hybridoma cell lines secreting the monoclonal antibody are another aspect of this invention.

Alternatively phage display technology could be utilized to select antibody genes with binding activities towards the polypeptide either from repertoires of PCR amplified v-genes of lymphocytes from humans screened for possessing anti-Fbp or from naive libraries (McCafferty, J. et al., Nature 348:552-554(1990), and Marks, J. et al., Biotechnology 10:779-783(1992)). The affinity of these antibodies can also be improved by chain shuffling (Clackson, T. et al., Nature 352:624-628(1991)).

The antibody should be screened again for high affinity to the polypeptide and/or fusion protein.

As mentioned above, a fragment of the final antibody may be prepared.

The antibody may be either intact antibody of M_r approx 150,000 or a derivative of it, for example a Fab fragment or a Fv fragment as described in Skerra, A and Pluckthun, A., Science 240:1038-1040 (1988). If two antigen binding domains are present each domain may be directed against a different epitope - termed 'bispecific' antibodies.

The antibody of the invention may be prepared by conventional means for example by established monoclonal antibody technology (Kohler, G. and Milstein, C. (supra, (1975) or using recombinant means e.g. combinatorial libraries, for example as described in Huse, W.D. *et al.*, Science 246:1275-1281 (1989).

5 Preferably the antibody is prepared by expression of a DNA polymer encoding said antibody in an appropriate expression system such as described above for the expression of polypeptides of the invention. The choice of vector for the expression system will be determined in part by the host, which may be a prokaryotic cell, such as *E. coli* (preferably strain B) or *Streptomyces sp.* or a eukaryotic cell, such as a mouse C127, mouse myeloma,
10 human HeLa, Chinese hamster ovary, filamentous or unicellular fungi or insect cell. The host may also be a transgenic animal or a transgenic plant (for example, as described in Hiatt, A. *et al.*, Nature 340:76-78(1989). Suitable vectors include plasmids, bacteriophages, cosmids and recombinant viruses, derived from, for example, baculoviruses and vaccinia.

 The Fab fragment may also be prepared from its parent monoclonal antibody by
15 enzyme treatment, for example using papain to cleave the Fab portion from the Fc portion.

 Preferably the antibody or derivative thereof is modified to make it less immunogenic in the patient. For example, if the patient is human the antibody may most preferably be 'humanized'; where the complementarity determining region(s) of the hybridoma-derived antibody has been transplanted into a human monoclonal antibody, for
20 example as described in Jones, P. *et al.*, Nature 321:522-525 (1986), or Tempest *et al.*, Biotechnology 9:266-273 (1991).

 The modification need not be restricted to one of 'humanization' ; other primate sequences (for example Newman, R. *et al.*, Biotechnology 10:1455-1460 (1992)) may also be used.

25 The humanized monoclonal antibody, or its fragment having binding activity, form a particular aspect of this invention.

This invention provides a method of screening compounds, for example, drugs to identify those which activate or preferably interfere with the proteins selected as targets herein, which method comprises measuring the activation or interference of the activity of the protein by a test compound or drug. For example if the protein selected has a catalytic activity, after suitable purification and formulation the activity of the enzyme can be followed by its ability to convert its natural substrates. By incorporating different chemically synthesized test compounds or natural products into such an assay of enzymatic activity one is able to detect those additives which compete with the natural substrate or otherwise inhibit enzymatic activity.

The invention also relates to inhibitors identified thereby.

The use of a polynucleotide of the invention in genetic immunization will preferably employ a suitable delivery method such as direct injection of plasmid DNA into muscles (Wolff *et al.*, Hum. Mol. Genet. 1:363 (1992); Manthorpe *et al.*, Hum. Gene Ther. 4:419 (1993)), delivery of DNA complexed with specific protein carriers (Wu *et al.*, J. Biol. Chem. 264:16985 (1989)), coprecipitation of DNA with calcium phosphate (Benvenisty & Reshef, Proc. Nat'l Acad. Sci. USA, 83:9551 (1986)), encapsulation of DNA in various forms of liposomes (Kaneda *et al.*, Science 243:375 (1989)), particle bombardment (Tang *et al.*, Nature 356:152 (1992)); Eisenbraun *et al.*, DNA Cell Biol. 12:791 (1993)) and *in vivo* infection using cloned retroviral vectors (Seeger *et al.*, Proc. Nat'l Acad. Sci. USA 81:5849 (1984)). Suitable promoters for muscle transfection include CMV, RSV, SRa, actin, MCK, alpha globin, adenovirus and dihydrofolate reductase.

In therapy or as a prophylactic, the active agent i.e., the polypeptide, polynucleotide or inhibitor of the invention, may be administered to a patient as an injectable composition, for example as a sterile aqueous dispersion, preferably isotonic.

Alternatively the composition may be formulated for topical application for example in the form of ointments, creams, lotions, eye ointments, eye drops, ear drops, mouthwash, impregnated dressings and sutures and aerosols, and may contain appropriate conventional additives, including, for example, preservatives, solvents to assist drug penetration, and emollients in ointments and creams. Such topical formulations may also contain compatible conventional carriers, for example cream or ointment bases, and ethanol or oleyl alcohol for lotions. Such carriers may constitute from about 1% to about 98% by weight of the formulation; more usually they will constitute up to about 80% by weight of the formulation.

For administration to human patients, it is expected that the daily dosage level of the active agent will be from 0.01 to 10 mg/kg, typically around 1 mg/kg. The physician in any event will determine the actual dosage which will be most suitable for an individual patient and will vary with the age, weight and response of the particular patient. The above dosages are exemplary of the average case. There can, of course, be individual instances where higher or lower dosage ranges are merited, and such are within the scope of this invention.

A vaccine composition is conveniently in injectable form. Conventional adjuvants may be employed to enhance the immune response.

A suitable unit dose for vaccination is 0.5-5 μ g/kg of antigen, and such dose is preferably administered 1-3 times and with an interval of 1-3 weeks.

Within the indicated dosage range, no adverse toxicological effects are expected with the compounds of the invention which would preclude their administration to suitable patients.

EXAMPLES

In order to facilitate understanding of the following non-limiting examples certain frequently occurring methods and/or terms will be described.

"Plasmids" are designated by a lower case p preceded and/or followed by capital letters and/or numbers. The starting plasmids herein are either commercially available, publicly available on an unrestricted basis, or can be constructed from available plasmids in accord with published procedures. In addition, equivalent plasmids to those described are known in the art and will be apparent to the ordinarily skilled artisan.

"Digestion" of DNA refers to catalytic cleavage of the DNA with a restriction enzyme that acts only at certain sequences in the DNA. The various restriction enzymes used herein are commercially available and their reaction conditions, cofactors and other requirements were used as would be known to the ordinarily skilled artisan. For analytical purposes, typically 1 μ g of plasmid or DNA fragment is used with about 2 units of enzyme in about 20 μ l of buffer solution. For the purpose of isolating DNA fragments for plasmid construction, typically 5 to 50 μ g of DNA are digested with 20 to 250 units of enzyme in a larger volume. Appropriate buffers and substrate amounts for particular restriction enzymes are specified by the manufacturer. Incubation times of about 1 hour at 37 C are ordinarily used, but may vary in accordance with the supplier's instructions. After digestion the reaction is electrophoresed directly on a polyacrylamide gel to isolate the desired fragment.

Size separation of the cleaved fragments is performed using 8 percent polyacrylamide gel described by Goeddel, D. *et al.*, (1980) *Nucleic Acids Res.*, 8:4057.

"Oligonucleotides" refers to either a single stranded polydeoxynucleotide or two complementary polydeoxynucleotide strands which may be chemically synthesized. Such synthetic oligonucleotides have no 5' phosphate and thus will not ligate to another oligonucleotide without adding a phosphate with an ATP in the presence of a kinase. A synthetic oligonucleotide will ligate to a fragment that has not been dephosphorylated.

"Ligation" refers to the process of forming phosphodiester bonds between two double stranded nucleic acid fragments (Maniatis, T., *et al.*, *supra*, p. 146). Unless otherwise provided, ligation may be accomplished using known buffers and conditions with 10 units to T4 DNA ligase ("ligase") per 0.5 µg of approximately equimolar amounts of the DNA fragments to be ligated.

Example 1

Isolation of DNA coding for a Novel Protein from *S. Aureus* WCUH 29

The polynucleotide having the DNA sequence given herein can be obtained from a library of clones of chromosomal DNA of *S.aureus* WCUH 29 in *E.coli*. Libraries may be prepared by routine methods, for example:

Methods 1 and 2.

Total cellular DNA is isolated from *Staphylococcus aureus* strain WCUH29 (NCIMB 40771) according to standard procedures and size-fractionated by either of two methods.

Method 1.

Total cellular DNA is mechanically sheared by passage through a needle in order to size-fractionate according to standard procedures. DNA fragments of up to 1 kbp in size are rendered blunt by treatment with exonuclease and DNA polymerase, and *EcoRI* linkers added. Fragments are ligated into the vector Lambda ZapII that has been cut with *EcoRI*, the library packaged by standard procedures and *E.coli* infected with the packaged library. The library is amplified by standard procedures.

Method 2.

Total cellular DNA is partially hydrolyzed with a combination of four restriction enzymes (*RsaI*, *PstI*, *AluI* and *Bsh1235I*) and size-fractionated according to standard procedures. *EcoRI* linkers are ligated to the DNA and the fragments then ligated into the vector Lambda ZapII that have been cut with *EcoRI*, the library packaged by standard

procedures, and *E.coli* infected with the packaged library. The library is amplified by standard procedures.

Example 2

The determination of expression during infection of a gene from *Staphylococcus aureus* WCUH29

Necrotic fatty tissue from a four day groin infection of *Staphylococcus aureus* WCUH29 in the mouse is efficiently disrupted and processed in the presence of chaotropic agents and RNAase inhibitor to provide a mixture of animal and bacterial RNA. The optimal conditions for disruption and processing to give stable preparations and high yields of bacterial RNA are followed by the use of hybridization to a radiolabelled oligonucleotide specific to *Staphylococcus aureus* 16S RNA on Northern blots. The RNase free, DNase free, DNA and protein free preparations of RNA obtained are suitable for Reverse Transcription PCR (RT-PCR) using unique primer pairs designed from the sequence of each gene of *Staphylococcus aureus* WCUH29.

a) Isolation of tissue infected with *Staphylococcus aureus* WCUH29 from a mouse animal model of infection

10 ml. volumes of sterile nutrient broth (No.2 Oxoid) are seeded with isolated, individual colonies of *Staphylococcus aureus* WCUH29 from an agar culture plate.

The cultures are incubated aerobically (static culture) at 37 degrees C for 16-20 hours. 4 week old mice (female, 18g-22g, strain MF1) are each infected by subcutaneous injection of 0.5ml. of this broth culture of *Staphylococcus aureus* WCUH29 (diluted in broth to approximately 10⁸ cfu/ml.) into the anterior, right lower quadrant (groin area). Mice should be monitored regularly during the first 24 hours after infection, then daily until termination of study. Animals with signs of systemic infection, i.e. lethargy, ruffled appearance, isolation from group, should be monitored closely and if signs progress to moribundancy, the animal should be culled immediately.

Visible external signs of lesion development will be seen 24-48h after infection. Examination of the abdomen of the animal will show the raised outline of the abscess beneath the skin. The localized lesion should remain in the right lower quadrant, but may occasionally spread to the left lower quadrant, and superiorly to the thorax. On occasions, the abscess may rupture through the overlying skin layers. The affected animal should be culled immediately and the tissues sampled if possible. Failure to cull the animal may result in the necrotic skin tissue overlying the abscess being sloughed off, exposing the abdominal muscle wall.

Approximately 96h after infection, animals are killed using carbon dioxide asphyxiation. To minimize delay between death and tissue processing /storage, mice should be killed individually rather than in groups. The dead animal is placed onto its back and the fur swabbed liberally with 70% alcohol. An initial incision using scissors is made through the skin of the abdominal left lower quadrant, travelling superiorly up to, then across the thorax. The incision is completed by cutting inferiorly to the abdominal lower right quadrant. Care should be taken not to penetrate the abdominal wall. Holding the skin flap with forceps, the skin is gently pulled away from the abdomen. The exposed abscess, which covers the peritoneal wall but generally does not penetrate the muscle sheet completely, is excised, taking care not to puncture the viscera.

The abscess/muscle sheet and other infected tissue, such as the necrotic pads of fatty tissue in the abdominal lower right and left quadrants may require cutting in sections, prior to flash-freezing in liquid nitrogen, thereby allowing easier storage in plastic collecting vials.

b) Isolation of *Staphylococcus aureus* WCUH29 RNA from infected tissue samples

4-6 infected tissue samples(each approx 0.5-0.7g) in 2ml screw-cap tubes are removed from -80°C storage into a dry ice ethanol bath. In a microbiological safety cabinet the samples are disrupted individually whilst the remaining samples are kept cold in the dry ice ethanol bath. To disrupt the bacteria within the tissue sample 1ml of TRIzol Reagent (Gibco BRL, Life Technologies) is added followed by enough 0.1mm zirconia/silica beads to almost fill the tube, the lid is replaced taking care not to get any beads into the screw thread so as to ensure a good seal and eliminate aerosol generation. The sample is then homogenized in a Mini-BeadBeater Type BX-4 (Biospec Products). Necrotic fatty tissue is treated for 100 seconds at 5000 rpm in order to achieve bacterial lysis. *In vivo* grown bacteria require longer treatment than *in vitro* grown *S.aureus* WCUH29 which are disrupted by a 30 second bead-beat.

After bead-beating the tubes are chilled on ice before opening in a fume-hood as heat generated during disruption may degrade the TRIzol and release cyanide.

200 microlitres of chloroform is then added and the tubes shaken by hand for 15 seconds to ensure complete mixing. After 2-3 minutes at room temperature the tubes are spun down at 12,000 x g, 4°C for 15 minutes and RNA extraction is then continued according to the method given by the manufacturers of TRIzol Reagent i.e.:- The aqueous phase, approx 0.6 ml, is transferred to a sterile eppendorf tube and 0.5 ml of isopropanol is added. After 10 minutes at room temperature the samples are spun at 12,000 x g, 4°C for 10

minutes. The supernatant is removed and discarded then the RNA pellet is washed with 1. ml 75% ethanol. A brief vortex is used to mix the sample before centrifuging at 7,500 x g, 4 °C for 5 minutes. The ethanol is removed and the RNA pellet dried under vacuum for no more than 5 minutes. Samples are then resuspended by repeated pipetting in 100
5 microlitres of DEPC treated water, followed by 5-10 minutes at 55 °C. Finally, after at least 1 minute on ice, 200 units of Rnasin (Promega) is added
RNA preparations are stored at -80 °C for up to one month. For longer term storage the RNA precipitate can be stored at the wash stage of the protocol in 75% ethanol for at least one year at -20 °C.

- 10 Quality of the RNA isolated is assessed by running samples on 1% agarose gels. 1x TBE gels stained with ethidium bromide are used to visualize total RNA yields. To demonstrate the isolation of bacterial RNA from the infected tissue 1 x MOPS, 2.2M formaldehyde gels are run and vacuum blotted to Hybond-N (Amersham). The blot is then hybridized with a ³²P labelled oligonucleotide probe specific to 16s rRNA of *S.aureus* (K.
15 Greisen, *et al.*, J. Clin. Microbiol. 32 335-351(1994)). An oligonucleotide of the sequence: 5'-gctcctaaaagggttactccaccggc-3' [SEQ ID NO:1166] is used as a probe. The size of the hybridizing band is compared to that of control RNA isolated from *in vitro* grown *S.aureus* WCUH29. Correct sized bacterial 16s rRNA bands can be detected in total RNA samples which show extensive degradation of the mammalian RNA when visualized on TBE gels.
20 **c) The removal of DNA from Staphylococcus aureus WCUH29 derived RNA**

DNA was removed from 73 microlitre samples of RNA by a 15 minute treatment on ice with 3 units of DNaseI, amplification grade (Gibco BRL, Life Technologies) in the buffer supplied with the addition of 200 units of Rnasin (Promega) in a final volume of 90 microlitres.

- 25 The DNase was inactivated and removed by treatment with TRIzol LS Reagent (Gibco BRL, Life Technologies) according to the manufacturers protocol.
DNase treated RNA was resuspended in 73 microlitres of DEPC treated water with the addition of Rnasin as described in Method 1.

d) The preparation of cDNA from RNA samples derived from infected tissue

- 30 10 microlitre samples of DNase treated RNA are reverse transcribed using a SuperScript Preamplification System for First Strand cDNA Synthesis kit (Gibco BRL, Life Technologies) according to the manufacturers instructions. 1 nanogram of random hexamers is used to prime each reaction. Controls without the addition of SuperScriptII reverse

transcriptase are also run. Both +/-RT samples are treated with RNaseH before proceeding to the PCR reaction

e) The use of PCR to determine the presence of a bacterial cDNA species

PCR reactions are set up on ice in 0.2ml tubes by adding the following

5 components:

45 microlitres PCR SUPERMIX (Gibco BRL, Life Technologies).

1 microlitre 50mM MgCl₂, to adjust final concentration to 2.5mM.

1 microlitre PCR primers (optimally 18-25 basepairs designed to possess similar annealing temperatures), each primer at 10mM initial concentration.

10 2 microlitres cDNA.

PCR reactions are run on a Perkin Elmer GeneAmp PCR System 9600 as follows:

5 minutes at 95 °C, then 50 cycles of 30 seconds each at 94 °C, 42 °C and 72 °C followed by 3 minutes at 72 °C and then a hold temperature of 4 °C.

15 10 microlitre aliquots are then run out on 1% 1 x TBE gels stained with ethidium bromide with PCR product sizes estimated by comparison to a 100 bp DNA Ladder (Gibco BRL, Life Technologies).

RT/PCR controls may include +/- reverse transcriptase reactions, 16s rRNA primers or DNA specific primer pairs designed to produce PCR products from non-transcribed *S.aureus* WCUH29 genomic sequences.

20 To test the efficiency of the primer pairs they are used in DNA PCR with WCUH29 total DNA. PCR reactions are set up and run as described above using approx. 1 microgram of DNA in place of the cDNA and 35 cycles of PCR rather than 50.

Primer pairs which fail to give the predicted sized product in either DNA PCR or RT/PCR (approx 20%) are PCR failures and as such are uninformative. Of those which give
25 the correct size product with DNA PCR three classes are distinguished in RT/PCR:

1. Genes which are not expressed *in vivo* reproducibly fail to give a product in RT/PCR.

2. Genes which are expressed *in vivo* reproducibly give the correct size product in RT/PCR and show a stronger signal in the +RT samples than in the -RT controls.

30 3. Genes which may be expressed *in vivo* give similar amounts of product in both +/-RT samples.

Table 1 sets forth certain information pertinent to cloning and identification of ORFs of the invention, such as reading frame direction (forward or reverse) on each polynucleotide in the Sequence Listing, a putative identity based on homology searches, the numbers of the first and last nucleic acid of the reading frame, and the number of codons in the ORF (including the stop codon). Using this information, expression vectors providing the selected protein are prepared and the protein is configured in an appropriate screen for the identification of anti-microbial agents. The SEQ ID NOS provided herewith for the amino acid (polypeptide) sequences are consecutively numbered to correlate with the information forth in Table 1, such that SEQ ID NO: 1 corresponds to the first sequence referenced in Table 1 and SEQ ID NO:623 corresponds to the last amino acid sequence referenced in Table 1. For the polynucleotide sequences SEQ ID NO:624 corresponds to the first polynucleotide sequence referenced in Table 1 and SEQ ID NO:1165 corresponds to the last sequence referenced Table 1.

Table 1

SEQ ID NO: __1__ ENCODED BY SEQ ID NO: __624__

ORF #	Start	End	Direction	Codon Length
1	64	396	F	111 codons

ORF translation from 64-396, direction F

Blastp and/or MPSearch Result:

Result: UNK

SEQ ID NO: __2__ ENCODED BY SEQ ID NO: __625__

ORF #	Start	End	Direction	Codon Length
1	127	375	F	83 codons

ORF translation from 127-375, direction F

Blastp and/or MPSearch Result:

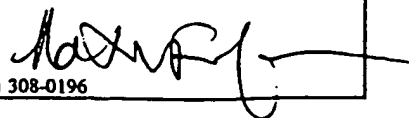
Result: UNK

WHAT IS CLAIMED IS:

1. An isolated polypeptide comprising an amino acid sequence selected from the group consisting of SEQ ID NOS: 8, 10, 14, 19, 25, 27, 28, 30, 32 and 36.
2. An isolated nucleic acid encoding one of the amino acid sequences of
5 Claim 1 and nucleic acid sequences capable of hybridizing therewith under stringent conditions.
3. Recombinant vectors comprising the nucleic acid sequences of Claim 2 and host cells transformed or transfected therewith.
4. A method of identifying an antimicrobial compound comprising
10 contacting candidate compounds with a polypeptide of Claim 1 and selecting those compounds capable of inhibiting the bioactivity of said polypeptide.
5. Antimicrobial compounds identified by the method of Claim 4.
6. An isolated polypeptide comprising an acid sequence selected from the group consisting of SEQ ID NO:1 through SEQ ID NO:623.
- 15 7. An isolated nucleic acid encoding one of the amino acid sequences of Claim 6 and nucleic acid sequences capable of hybridizing therewith under stringent conditions.
8. An isolated polynucleotide comprising a nucleic acid sequence selected from the group consisting of SEQ ID NOS: 631, 633, 636, 641, 647, 649,
20 650, 652, 654 and 658.
9. An isolated polynucleotide comprising a nucleic acid sequence selected from the group consisting of SEQ ID NO:624 through SEQ ID NO:1165.
10. An antibody against the polypeptide of claim 6.

INTERNATIONAL SEARCH REPORT

International application No. -
PCT/US97/02318

A. CLASSIFICATION OF SUBJECT MATTER IPC(6) : C07K 1/00, 16/00; C07H 21/04; C12N 15/00 US CL : 530/350, 387.9; 435/320.1, 325; 536/23.7 According to International Patent Classification (IPC) or to both national classification and IPC		
B. FIELDS SEARCHED Minimum documentation searched (classification system followed by classification symbols) U.S. : 530/350, 387.9; 435/320.1, 325; 536/23.7 Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched Electronic data base consulted during the international search (name of data base and, where practicable, search terms used) APS, MEDLINE, BIOSIS, EMBASE, CA, WPIDS terms: antimicrobial, antibiotic, vector, antibody		
C. DOCUMENTS CONSIDERED TO BE RELEVANT		
Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	US 4,431,739 A (RIGGS ET AL) 14 February 1984 (14/02/84), see entire document.	1-10
A	EP 0,233,715 A2 (BEECHAM GROUP PLC) 26 August 1987 (26/08/87), see entire document.	1-10
A	SEVIER et al. Monoclonal Antibodies in Clinical Immunology. Clinical Chemistry. November 1981, Vol. 27, No. 11, pages 1797-1806, see entire document.	10
A	AGA et al. Isolation and Identification of Antimicrobial Compounds in Brazilian Propolis. Biosci. Biotech. Biochem. May 1994, Vol. 58, No. 5, pages 945-946, see entire document.	4-5
<input type="checkbox"/> Further documents are listed in the continuation of Box C. <input type="checkbox"/> See patent family annex.		
* Special categories of cited documents: "A" document defining the general state of the art which is not considered to be of particular relevance "E" earlier documents published on or after the international filing date "L" documents which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reasons (as specified) "O" documents referring to an oral disclosure, use, exhibition or other means "P" documents published prior to the international filing date but later than the priority date claimed "T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention "X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone "Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art "Z" document member of the same patent family		
Date of the actual completion of the international search 30 APRIL 1997		Date of mailing of the international search report 24 JUN 1997
Name and mailing address of the ISA/US Commissioner of Patents and Trademarks Box PCT Washington, D.C. 20231 Facsimile No. (703) 305-3230		Authorized officer MARK NAVARRO  Telephone No. (703) 308-0196

INTERNATIONAL SEARCH REPORT

International application No. -
PCT/US97/02318

Box I Observations where certain claims were found unsearchable (Continuation of item 1 of first sheet)

This international report has not been established in respect of certain claims under Article 17(2)(a) for the following reasons:

1. ☐ Claims Nos.:
because they relate to subject matter not required to be searched by this Authority, namely:
2. ☐ Claims Nos.:
because they relate to parts of the international application that do not comply with the prescribed requirements to such an extent that no meaningful international search can be carried out, specifically:
3. ☐ Claims Nos.:
because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a).

Box II Observations where unity of invention is lacking (Continuation of item 2 of first sheet)

This International Searching Authority found multiple inventions in this international application, as follows:

Please See Extra Sheet.

1. ☐ As all required additional search fees were timely paid by the applicant, this international search report covers all searchable claims.
2. ☐ As all searchable claims could be searched without effort justifying an additional fee, this Authority did not invite payment of any additional fee.
3. ☒ As only some of the required additional search fees were timely paid by the applicant, this international search report covers only those claims for which fees were paid, specifically claims Nos.:
1-10 and the species recited in claims 1 and 8.
4. ☐ No required additional search fees were timely paid by the applicant. Consequently, this international search report is restricted to the invention first mentioned in the claims; it is covered by claims Nos.:

Remark on Protest

- ☐ The additional search fees were accompanied by the applicant's protest.
☐ No protest accompanied the payment of additional search fees.

INTERNATIONAL SEARCH REPORT

International application No. -
PCT/US97/02318

BOX II. OBSERVATIONS WHERE UNITY OF INVENTION WAS LACKING

This ISA found multiple inventions as follows:

This application contains the following inventions or groups of inventions which are not so linked as to form a single inventive concept under PCT Rule 13.1. In order for all inventions to be searched, the appropriate additional search fees must be paid.

Group I, claims 1-4, and 6-10, drawn to DNA, proteins, and antibodies.

Group II, claims 5, drawn to antimicrobial compounds.

The inventions listed as Groups I and II do not relate to a single inventive concept under PCT Rule 13.1 because, under PCT Rule 13.2, they lack the same or corresponding special technical features for the following reasons: The special technical feature of Group I, drawn to polypeptides and DNA selected from SEQ ID NOS 8, 10, 14, 19, 25, 27, 28, 30, 32, 36, and 631, 633, 636, 641, 647, 649, 650, 652, 654, and 658 is that they are polypeptides which are required for bioactivity while the special technical feature of the Group II invention is that the compounds identified are capable of inhibiting the bioactivity of said polypeptides.

This application contains claims directed to more than one species of the generic invention. These species are deemed to lack Unity of Invention because they are not so linked to form a single inventive concept under PCT Rule 13.1. In order for more than one species to be searched, the appropriate additional search fees must be paid. The species are as follows:

Species 1, SEQ ID NOS 8, 10, 14, 19, 25, 27, 28, 30, 32, 36, and 631, 633, 636, 641, 647, 649, 650, 652, 654, and 658.

Species 2-154, The remaining SEQ ID NOS in groups of 4 peptides and their corresponding DNA sequences.

The species listed above do not relate to a single inventive concept under PCT Rule 13.1 because, under PCT Rule 13.2, the species lack the same or corresponding special technical features for the following reasons: Each protein is a structurally distinct chemical compound and are unrelated to one another. Up to 10 independent and distinct peptides will be examined in a single application, and then up to 4 such additional sequences will constitute as an additional group.